Linear Algebra and its Applications 436 (2012) 1763–1769

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

# Hölder type matrix inequalities of Pate, Blakley, and Roy extended to the inner product of Frobenius

Thomas H. Pate

*Mathematics Department, Auburn University, Auburn, AL 36849, USA*

ARTICLE INFO

*Article history:*
Received 27 May 2011
Accepted 7 September 2011
Available online 5 November 2011

Submitted by R.A. Brualdi

*AMS classification:*
15A45
15A42
15A60
15A63

*Keywords:*
Symmetric products
Multilinear identities
Tensor inequalities

ABSTRACT

Suppose $m, n$, and $k$ are positive integers, and let $\langle\cdot,\cdot\rangle$ be the standard inner product on the spaces $\mathbb{R}^p$, $p > 0$. Recently Pate has shown that if $D$ is an $m \times n$ non-negative real matrix, and $u$ and $v$ are non-negative unit vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, then

$$\langle (DD^t)^k Du, v\rangle \;\geq\; \langle Du, v\rangle^{2k+1},$$

with equality if and only if $\langle (DD^t)^k Du, v\rangle = 0$, or there exists $\alpha > 0$ such that $Du = \alpha v$ and $D^t v = \alpha u$. This extends to non-symmetric non-square matrices a 1965 result of Blakley and Roy, and resolves a special case of a graph theoretic inequality conjectured by Sidorenko. We generalize the above, obtaining pure matrix inequalities involving the Frobenius inner product, $\langle\cdot,\cdot\rangle_f$. In particular, we show that if $k$ is a positive integer, and $D$, $X$, and $Y$ are non-negative matrices that are $m \times n$, $n \times p$, and $m \times p$, respectively, then

$$\left(\sum_{i=1}^{p} \|x_i\|\,\|y_i\|\right)^{2k} \langle D(D^tD)^k X, Y\rangle_f \;\geq\; (\langle DX, Y\rangle_f)^{2k+1},$$

where $X$ has columns $x_1, x_2, \ldots, x_p$, $Y$ has columns $y_1, y_2, \ldots, y_p$, and $\|\cdot\|$ is the 2-norm. Necessary and sufficient conditions for equality are also given.



## 1. Introduction

We let $\mathbb{R}^{m\times n}$ and $\mathbb{C}^{m\times n}$ denote, respectively, the set of all $m \times n$ real matrices, and the set of all $m \times n$ complex matrices provided with the Frobenius inner product, $\langle\cdot,\cdot\rangle_f$, defined by $\langle A, B\rangle_f =$

*E-mail address:* pate_tom@bellsouth.net

0024-3795/$ - see front matter 
doi:10.1016/j.laa.2011.09.019

$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\overline{b_{ij}} = \mathrm{trace}(AB^*)$ for all $A = [a_{ij}]$, and $B = [b_{ij}]$ in $\mathbb{C}^{m\times n}$. If $B \in \mathbb{C}^{m\times n}$, then $B^t$ denotes the transpose of $B$, and $B^*$ denotes the conjugate transpose of $B$. The set of all $m \times n$ non-negative real matrices, a primary focus, is denoted by $\mathbb{R}_+^{m\times n}$, and we let $\mathbb{R}_+^n$, essentially $\mathbb{R}_+^{n\times 1}$, denote the set of all non-negative real vectors of length $n$. The notations $\mathbb{R}^n$ and $\mathbb{C}^n$ will have their usual meanings, and we equip these vector spaces with standard inner product, denoted by $\langle\cdot,\cdot\rangle$; thus, if $x = (x_1, x_2, \ldots, x_n)^t$, and $y = (y_1, y_2, \ldots, y_n)^t$ are in $\mathbb{R}^n$, then $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$, while if $x$ and $y$ are in $\mathbb{C}^n$, then $\langle x, y\rangle = \sum_{i=1}^{n} x_i \bar{y}_i$. The norm associated with $\langle\cdot,\cdot\rangle$, the 2-norm, is denoted by $\|\cdot\|$, and $\|\cdot\|_f$ is the norm associated with $\langle\cdot,\cdot\rangle_f$.

In the 1965 paper "A Hölder Type Inequality for Symmetric Matrices with Non-negative Entries" Blakely and Roy [1] proved

**Theorem 1.** *Suppose n and k are positive integers. If D is a symmetric member of* $\mathbb{R}_+^{n\times n}$, *then*

$$\langle D^k u, u\rangle \geq \langle Du, u\rangle^k, \quad \forall u \in \mathbb{R}_+^n \ \ \textit{such that} \ \ \|u\| = 1, \tag{1}$$

*with equality, when* $k \geq 2$, *if and only if* $\langle D^k u, u\rangle = 0$, *or there exists* $\alpha > 0$ *such that* $Du = \alpha u$.

In [4] this author published the following considerably stronger result wherein it is not assumed that $D$ is symmetric or even square. Moreover, the single unit vector $u$ appearing in Theorem 1 has been replaced by a pair of vectors $u$ and $v$.

**Theorem 2.** *If m, n, and k are positive integers, and D is a member of* $\mathbb{R}_+^{m\times n}$, *then*

$$(\|u\|\,\|v\|)^{2k}\langle (DD^t)^k Du, v\rangle \geq \langle Du, v\rangle^{2k+1}, \tag{2}$$

*for all* $u \in \mathbb{R}_+^n$, *and* $v \in \mathbb{R}_+^m$, *with equality if and only if* $\langle (DD^t)^k Du, v\rangle = 0$, *or there exists positive numbers* $\xi$ *and* $\alpha$ *such that* $Du = (\xi\alpha)\,v$, *and* $D^t v = (\xi/\alpha)\,u$.

The number $\alpha$ referenced in the necessary and sufficient condition for equality in Theorem 2 simply must be $\|u\|/\|v\|$, as the condition $Du = (\xi\alpha)v$ implies that $\langle Du, v\rangle = (\xi\alpha)\|v\|^2$, and the condition $D^t v = (\xi/\alpha)u$ implies that $\langle Du, v\rangle = \langle D^t v, u\rangle = (\xi/\alpha)\|u\|^2$, and these two together imply that $(\xi\alpha)\|v\|^2 = (\xi/\alpha)\|u\|^2$, which, since $\xi > 0$, is equvalent to the statement that $\alpha = \|u\|/\|v\|$.

When it is assumed that $\|u\| = \|v\| = 1$, Theorem 2 becomes

**Theorem 3.** *If m, n, and k are positive integers, and D is a member of* $\mathbb{R}_+^{m\times n}$, *then*

$$\langle (DD^t)^k Du, v\rangle \geq \langle Du, v\rangle^{2k+1}, \tag{3}$$

*for all unit* $u \in \mathbb{R}_+^n$, *and all unit* $v \in \mathbb{R}_+^m$, *with equality if and only if* $\langle (DD^t)^k Du, v\rangle = 0$, *or there exists a positive number* $\xi$ *such that* $Du = \xi\, v$ *and* $D^t v = \xi\, u$.

## 2. New results

As noted above our results involve the Frobenius inner product $\langle\cdot,\cdot\rangle_f$. Recall that if $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times p}$, and $C \in \mathbb{C}^{m\times p}$, then

$$\langle AB, C\rangle_f = \langle A, CB^*\rangle_f = \langle B, A^*C\rangle_f. \tag{4}$$

This is because $\langle AB, C\rangle_f = \mathrm{trace}(ABC^*) = \mathrm{trace}(A(CB^*)^*) = \langle A, CB^*\rangle_f$, and $\langle AB, C\rangle_f = \mathrm{trace}(ABC^*) = \mathrm{trace}(BC^*A) = \mathrm{trace}(B(A^*C)^*) = \langle B, A^*C\rangle_f$. Employing (4) we may rewrite (3) as

$$\langle (DD^t)^k D, vu^t\rangle_f \geq (\langle D, vu^t\rangle_f)^{2k+1}. \tag{5}$$

Mindful of inequality (2), expression (5) leads us to search for an inequality of the form

$$H(A)\langle (DD^t)^k D, A\rangle_f \;\geq\; (\langle D, A\rangle_f)^{2k+1} \tag{6}$$

that would hold for all $A \in \mathbb{R}_+^{m\times n}$, where $H(\cdot)$ is a function from $\mathbb{R}^{m\times n}$ to $[0,\infty)$. Actually, we obtain a result somewhat more general than (6).

If $x_1, x_2, \ldots, x_p$ are in $\mathbb{C}^n$, then we let $[x_1, x_2, \ldots, x_p]$ denote the $n\times p$ matrix whose $i$th column vector is $x_i$ for each $i \in \{1, 2, \ldots, p\}$. If $X = [x_1, x_2, \ldots, x_p]$, and $Y = [y_1, y_2, \ldots, y_p]$, where each $y_i \in \mathbb{C}^m$, then $\mathcal{F}(X, Y)$ denotes $\sum_{i=1}^p \|x_i\|\,\|y_i\|$. The following lemma contains the bulk of our main result. It differs from Theorem 4, our main result, only because we assume that neither $X$ nor $Y$ has a column of zeroes.

**Lemma 1.** *Suppose $m$, $n$, $k$ and $p$ are positive integers, and $D\in\mathbb{R}_+^{m\times n}$. If $X\in\mathbb{R}_+^{n\times p}$, $Y\in\mathbb{R}_+^{m\times p}$, and neither $X$ nor $Y$ has a column of zeroes, then*

$$(\mathcal{F}(X, Y))^{2k}\langle D(D^tD)^k X, Y\rangle_f \;\geq\; (\langle DX, Y\rangle_f)^{2k+1}, \tag{7}$$

*with equality if and only if $\langle D(D^tD)^k X, Y\rangle_f$ is 0, or there exists a positive number $\xi$, and a $p\times p$ diagonal matrix $\Lambda$ with positive diagonal such that $DX = \xi\, Y\Lambda$, and $D^tY = \xi X\Lambda^{-1}$.*

**Proof.** We first prove the inequality. Let $X = [x_1, x_2, \ldots, x_p]$, and $Y = [y_1, y_2, \ldots, y_p]$, where each $x_i \in \mathbb{R}_+^n$, and each $y_i \in \mathbb{R}_+^m$. Let $\Gamma = D(D^tD)^k = (DD^t)^k D$, and note that

$$\langle\, \Gamma X, Y\,\rangle_f \;=\; \langle\, [\Gamma x_1, \Gamma x_2, \ldots, \Gamma x_p], [y_1, y_2, \ldots, y_p]\,\rangle_f \;=\; \sum_{i=1}^p \langle \Gamma x_i, y_i\rangle. \tag{8}$$

Since no $x_i$ or $y_i$ is 0, $\mathcal{F}(X, Y)$, which we denote by $c$, must be positive. Normalizing, we set $\tilde{x}_i = x_i/\|x_i\|$, $\tilde{y}_i = y_i/\|y_i\|$, and $\rho_i = \|x_i\|\,\|y_i\|/c$ for all $i$, $1\leq i\leq p$. Then, each $\rho_i > 0$, and $\sum_{i=1}^p \rho_i = 1$; moreover, $(1/c)\langle\Gamma x_i, y_i\rangle = \rho_i\langle\Gamma\tilde{x}_i, \tilde{y}_i\rangle$. Mindful of these details, we combine (8) and Theorem 3 to obtain that

$$(1/c)\langle\Gamma X, Y\rangle_f \;=\; (1/c)\sum_{i=1}^p\langle\Gamma x_i, y_i\rangle \;=\; \sum_{i=1}^p \rho_i\langle\Gamma\tilde{x}_i, \tilde{y}_i\rangle \;\geq\; \sum_{i=1}^p \rho_i(\langle D\tilde{x}_i, \tilde{y}_i\rangle)^{2k+1}. \tag{9}$$

Since $k\geq 1$, the function $z\mapsto z^{2k+1}$ is strictly convex on $[0,\infty)$; so, by Theorem 90 of [3], essentially Jensen's inequality, we obtain that

$$\begin{aligned}\sum_{i=1}^p \rho_i(\langle D\tilde{x}_i, \tilde{y}_i\rangle)^{2k+1} &\geq \left(\sum_{i=1}^p \rho_i\langle D\tilde{x}_i, \tilde{y}_i\rangle\right)^{2k+1} = \left(\sum_{i=1}^p (\|x_i\|\,\|y_i\|c^{-1})\langle D\tilde{x}_i, \tilde{y}_i\rangle\right)^{2k+1}\\ &= (1/c)^{2k+1}\left(\sum_{i=1}^p\langle Dx_i, y_i\rangle\right)^{2k+1} = (1/c)^{2k+1}(\langle DX, Y\rangle_f)^{2k+1}.\end{aligned} \tag{10}$$

In conjunction, (9) and (10) say that $(1/c)\langle\Gamma X, Y\rangle_f \geq (1/c)^{2k+1}(\langle DX, Y\rangle_f)^{2k+1}$, which is equivalent to

$$(\mathcal{F}(X, Y))^{2k}\langle D(D^tD)^k X, Y\rangle_f \;\geq\; (\langle DX, Y\rangle_f)^{2k+1}. \tag{11}$$

This completes the proof of the inequality (7).

To verify that the stated conditions are necessary and sufficient for equality we assume that neither $X$ nor $Y$ has a column of zeroes. We verify sufficiency first. If $\langle\Gamma X, Y\rangle_f = 0$, then, since we know that $\langle DX, Y\rangle_f \geq 0$, the inequality (7), which we have just finished proving, implies that

$$0 = (\mathcal{F}(X, Y))^{2k} \langle D(D^tD)^kX, Y\rangle_f \;\geq\; (\langle DX, Y\rangle_f)^{2k+1} \geq 0,$$

which, in turn, implies that $\langle DX, Y\rangle_f = 0$. Thus, if $\langle \Gamma X, Y\rangle_f = 0$, then both sides of (7) reduce to 0. If $\langle \Gamma X, Y\rangle_f \neq 0$, and there exists a positive number $\xi$, and a positive $p \times p$ diagonal matrix $\Lambda$ such that $DX = \xi Y\Lambda$ and $D^tY = \xi X\Lambda^{-1}$, then $(D^tD)X = \xi(D^tY)\Lambda = \xi^2X(\Lambda^{-1}\Lambda) = \xi^2X$; thus, $(D^tD)^kX = \xi^{2k}X$, and $D(D^tD)^kX = \xi^{2k}DX = \xi^{2k+1}Y\Lambda$. Let $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$. We need expressions for the $\lambda_i$. For each $i$ we have $Dx_i = \xi\lambda_iy_i$ and $D^ty_i = \xi\lambda_i^{-1}x_i$; thus, $\langle Dx_i, y_i\rangle = \xi\lambda_i\|y_i\|^2$, and $\langle x_i, D^ty_i\rangle = \langle Dx_i, y_i\rangle = \xi\lambda_i^{-1}\|x_i\|^2$. Consequently, $\xi\lambda_i\|y_i\|^2 = \xi\lambda_i^{-1}\|x_i\|^2$, which, since $\xi > 0$, tells us that $\lambda_i = \|x_i\|/\|y_i\|$ for each $i$. Let $c$ denote $\mathcal{F}(X, Y)$ as above, and note that $\langle Y\Lambda, Y\rangle_f = \sum_{i=1}^p \lambda_i\|y_i\|^2 = \sum_{i=1}^p \|x_i\|\,\|y_i\| = c$. Since $D(D^tD)^kX = \xi^{2k+1}Y\Lambda$, the left side of (7) is

$$(\mathcal{F}(X, Y))^{2k}\langle D(D^tD)^kX, Y\rangle_f = c^{2k}\langle D(D^tD)^kX, Y\rangle_f = c^{2k}\xi^{2k+1}\langle Y\Lambda, Y\rangle_f = c^{2k+1}\xi^{2k+1}. \tag{12}$$

On the other hand, $\langle DX, Y\rangle_f = \xi\langle Y\Lambda, Y\rangle_f = c\xi$, so the right side of (7) is also $c^{2k+1}\xi^{2k+1}$. The stated conditions are therefore sufficient for equality in (7).

To establish necessity we assume that (7) is an equality, and note that this implies that both (9) and (10) must reduce to equality. But each $\rho_i > 0$, so equality in (9) implies that $\langle \Gamma\tilde{x}_i, \tilde{y}_i\rangle = (\langle D\tilde{x}_i, \tilde{y}_i\rangle)^{2k+1}$ for each $i$ such that $1 \leq i \leq p$. If there exists one integer $j$ such that $1 \leq j \leq p$, and $\langle \Gamma\tilde{x}_j, \tilde{y}_j\rangle = 0$, then the corresponding term $\langle D\tilde{x}_j, \tilde{y}_j\rangle$ must be 0, which, because of the conditions for equality in Jensen's inequality [3, Theorem 95], and because (10) is an equality, implies that all of the $\langle D\tilde{x}_i, \tilde{y}_i\rangle$ must be zero. But then $0 = \sum_{i=1}^p \rho_i\langle D\tilde{x}_i, \tilde{y}_i\rangle = (1/c)\langle \Gamma X, Y\rangle_f$; that is, $\langle \Gamma X, Y\rangle_f = 0$. Thus, if some one of the $\langle \Gamma\tilde{x}_j, \tilde{y}_j\rangle$ is 0, then all are 0, and $\langle \Gamma X, Y\rangle_f = 0$. We assume therefore that there is no $j$ such that $\langle \Gamma\tilde{x}_j, \tilde{y}_j\rangle$ is 0. Then, because of the equality assumption in (7) no one of the $\langle D\tilde{x}_i, \tilde{y}_i\rangle$ is 0. On account of the condition for equality in Theorem 3 we know that there exists positive numbers $\xi_1, \xi_2, \ldots, \xi_p$ such that for each $i$ such that $1 \leq i \leq p$, we have both $D\tilde{x}_i = \xi_i\tilde{y}_i$, and $D^t\tilde{y}_i = \xi_i\tilde{x}_i$. For each $i$ let $\alpha_i$ denote $\|x_i\|/\|y_i\|$. Then, each $\alpha_i$ is positive, and we have

$$Dx_i = \alpha_i\xi_iy_i\,, \qquad \text{and} \qquad D^ty_i = \alpha_i^{-1}\xi_ix_i \qquad \forall i \in \{1, 2, \ldots, p\}. \tag{13}$$

From (13) we deduce that $\langle Dx_i, y_i\rangle = \xi_i\|x_i\|\,\|y_i\|$ for each $i$. Alternately, $\xi_i = \langle D\tilde{x}_i, \tilde{y}_i\rangle$ for each $i$. But, as before, equality in (10) implies that the $\langle D\tilde{x}_i, \tilde{y}_i\rangle$ are all the same; that is, we have $\xi_1 = \xi_2 = \cdots = \xi_p$. Letting $\xi$ denote the common value of the $\xi_i$, we may restate (13) as

$$Dx_i = \alpha_i\xi y_i\,, \qquad \text{and} \qquad D^ty_i = \alpha_i^{-1}\xi x_i \qquad \forall i \in \{1, 2, \ldots, p\}, \tag{14}$$

which, if we let $\Lambda$ denote the diagonal matrix $\mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_p)$, may be restated as

$$DX = \xi Y\Lambda, \qquad \text{and} \qquad D^tY = \xi X\Lambda^{-1}. \tag{15}$$

This completes the proof of Lemma 1. $\square$

To obtain Theorem 4, our main result, we will strengthen Lemma 1 by eliminating the assumption that neither $X$ nor $Y$ has a column of zeroes. By continuity, inequality (7) is still true when $X$ or $Y$ has columns of zeroes. It is the necessary and sufficient condition for equality that requires adjustment. To state Theorem 4 in such a way that it obviously reduces to Lemma 1 when neither $X$ nor $Y$ has a column of zeroes, we associate with each $X = [x_1, x_2, \ldots, x_p] \in \mathbb{C}^{m\times p}$ and $Y = [y_1, y_2, \ldots, y_p] \in \mathbb{C}^{n\times p}$ a $p\times p$ diagonal (0,1)-matrix $\mathcal{P}_{X,Y}$ such that $(\mathcal{P}_{X,Y})_{ii} = 1$ if and only if neither $x_i$ nor $y_i$ is 0. If $x_i \neq 0$ and $y_i \neq 0$ for all $i \in \{1, 2, \ldots, p\}$, then $\mathcal{P}_{X,Y} = \mathrm{I}_p$, the $p\times p$ identity matrix, and Theorem 4 reduces to Lemma 1.

**Theorem 4.** *Suppose $m$, $n$, $k$ and $p$ are positive integers, and $D \in \mathbb{R}_+^{m\times n}$. If $X \in \mathbb{R}_+^{n\times p}$, and $Y \in \mathbb{R}_+^{m\times p}$, then*

$$(\mathcal{F}(X, Y))^{2k} \langle D(D^tD)^kX, Y\rangle_f \;\geq\; (\langle DX, Y\rangle_f)^{2k+1}, \tag{16}$$

*with equality if and only if* $\langle D(D^tD)^kX, Y\rangle_f$ *is* 0*, or there exists a positive number* $\xi$*, and a* $p \times p$ *diagonal matrix* $\Lambda$ *with positive diagonal such that* $DX\mathcal{P} = \xi Y\mathcal{P}\Lambda$*, and* $D^tY\mathcal{P} = \xi X\mathcal{P}\Lambda^{-1}$*, where* $\mathcal{P} = \mathcal{P}_{X,Y}$*.*

**Proof.** As noted above, Lemma 1 and continuity imply that (16) is true when either $X$ or $Y$ has columns of zeroes. We turn to the necessary and sufficient condition for equality, considering sufficiency first. Clearly, we have equality in (16) when $\langle D(D^tD)^kX, Y\rangle_f = 0$, since both sides of (16) reduce to zero in this case. This follows from (16) which has already been established. So, we assume that $DX\mathcal{P} = \xi Y\mathcal{P}\Lambda$, and $D^tY\mathcal{P} = \xi X\mathcal{P}\Lambda^{-1}$ where $\xi > 0$, and $\Lambda$ is a $p \times p$ diagonal matrix with positive diagonal. A key point is that if $B$ is any $m \times n$ matrix whatever, then $\langle BX\mathcal{P}, Y\mathcal{P}\rangle_f = \langle BX, Y\rangle_f$. Now, we have $D^t(DX\mathcal{P}) = D^t(\xi Y\mathcal{P}\Lambda) = \xi(D^tY\mathcal{P})\Lambda = \xi(\xi X\mathcal{P}\Lambda^{-1})\Lambda = \xi^2 X\mathcal{P}$; thus, $(D^tD)^k(X\mathcal{P}) = \xi^{2k}X\mathcal{P}$, and $D(D^tD)^k(X\mathcal{P}) = \xi^{2k}DX\mathcal{P} = \xi^{2k+1}Y\mathcal{P}\Lambda$. This implies that

$$\langle D(D^tD)^kX, Y\rangle_f = \langle D(D^tD)^kX\mathcal{P}, Y\mathcal{P}\rangle_f = \xi^{2k+1}\langle Y\mathcal{P}\Lambda, Y\mathcal{P}\rangle_f. \tag{17}$$

Let $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$, let $J$ denote the set of all $i \in \{1, 2, \ldots, p\}$ such that neither $x_i$ nor $y_i$ is zero, and let $c = \mathcal{F}(X, Y)$ as above. Then, for $i \in J$ we have $\lambda_i = \|x_i\|/\|y_i\|$, so

$$\begin{aligned}\langle Y\mathcal{P}\Lambda, Y\mathcal{P}\rangle_f &= \sum_{i\in J}\lambda_i\|y_i\|^2 = \sum_{i\in J}(\|x_i\|/\|y_i\|)\|y_i\|^2\\ &= \sum_{i\in J}\|x_i\|\,\|y_i\| = \sum_{i=1}^{p}\|x_i\|\,\|y_i\| = \mathcal{F}(X, Y) = c.\end{aligned} \tag{18}$$

On account of (17) and (18) we have $\langle D(D^tD)^kX, Y\rangle_f = \xi^{2k+1}c$; thus $(\mathcal{F}(X, Y))^{2k}\langle D(D^tD)^kX, Y\rangle_f$, which is the left side of (16), is $c^{2k+1}\xi^{2k+1}$. On the other hand,

$$\langle DX, Y\rangle_f = \langle DX\mathcal{P}, Y\mathcal{P}\rangle_f = \xi\langle Y\mathcal{P}\Lambda, Y\mathcal{P}\rangle_f = \xi c, \tag{19}$$

so, the right side of (16), namely $(\langle DX, Y\rangle_f)^{2k+1}$, is also $\xi^{2k+1}c^{2k+1}$. The stated conditions are therefore sufficient for equality.

Now, assume that (16) is an equality. We will use Lemma 1. As above, let $J$ denote the set of all $i \in \{1, 2, \ldots, p\}$ such that neither $x_i$ nor $y_i$ is zero. If $J$ is empty, then $\langle D(D^tD)^kX, Y\rangle_f = 0$, so we assume that $J$ is non-empty. Let $q = |J|$, and let $\tilde{X}$ denote the $n \times q$ matrix obtained from $X$ by deleting columns of $X$ corresponding to integers $j \in \{1, 2, \ldots, p\} \setminus J$. In other words, we delete column $j$ from $X$ if either $x_i$ or $y_i$ is the zero vector. Define the $m \times q$ matrix $\tilde{Y}$ correspondingly. Applying Lemma 1 we deduce that

$$(\mathcal{F}(\tilde{X}, \tilde{Y}))^{2k}\langle D(D^tD)^k\tilde{X}, \tilde{Y}\rangle_f \geq (\langle D\tilde{X}, \tilde{Y}\rangle_f)^{2k+1}. \tag{20}$$

Because of the way $\tilde{X}$ and $\tilde{Y}$ are obtained from $X$ and $Y$ it is clear that

$$\mathcal{F}(\tilde{X}, \tilde{Y}) = \mathcal{F}(X, Y), \quad \langle D(D^tD)^k\tilde{X}, \tilde{Y}\rangle_f = \langle D(D^tD)^kX, Y\rangle_f, \quad \text{and} \quad \langle D\tilde{X}, \tilde{Y}\rangle_f = \langle DX, Y\rangle_f. \tag{21}$$

Therefore, equality in (16) implies that (20) is an equality, which, on account of Lemma 1, implies that $\langle D(D^tD)^k\tilde{X}, \tilde{Y}\rangle_f = 0$, or there exists a positive number $\xi$, and a $q \times q$ diagonal matrix $\tilde{\Lambda}$ with positive diagonal such that

$$D\tilde{X} = \xi\tilde{Y}\tilde{\Lambda}, \qquad \text{and} \qquad D^t\tilde{Y} = \xi\tilde{X}\tilde{\Lambda}^{-1}. \tag{22}$$

In the former case we must, on account of (21), have $\langle D(D^tD)^kX, Y\rangle_f = 0$. In the latter case we must have $(\tilde{\Lambda})_{ii} = \|\tilde{x}_i\|/\|\tilde{y}_i\|$ for each $i \in \{1, 2, \ldots, q\}$. Let $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$, where $\lambda_i = \|x_i\|/\|y_i\|$

if neither $x_i$ nor $y_i$ is 0, and $\lambda_{ii} = 1$ otherwise. To obtain $\tilde{\Lambda}$ from $\Lambda$ we simply delete the columns corresponding to the members of $\{1, 2, \ldots, p\} \backslash J$. But this is exactly how we obtain $\tilde{X}$ from $X$, and $\tilde{Y}$ from $Y$. Moreover, $X\mathcal{P}$ is obtained from $X$ by replacing all columns corresponding to members of $\{1, 2, \ldots, p\} \backslash J$ with columns of zeroes, and $Y\mathcal{P}$ is obtained from $Y$ in the same way. Therefore, (22) is equivalent to

$$DX\mathcal{P} = \xi Y\mathcal{P}\Lambda, \qquad \text{and} \qquad D^t Y\mathcal{P} = \xi X\mathcal{P}\Lambda^{-1}. \tag{23}$$

In essence, equality in (16) places no constraint upon columns of $X$ and $Y$ consisting entirely of zeroes. This is because such columns contribute nothing to the three major terms appearing in (16). This completes the proof of Theorem 4. □

Theorem 4 is a true extension of Theorem 2, as it reduces to Theorem 2 when $p = 1$. By specializing Theorem 4 we obtain other interesting results. For example, if we set $p = n$, and let $X$ be $\mathrm{I}_n$, then we obtain the following.

**Theorem 5.** *Suppose $m$, $n$, and $k$ are positive integers, and $D \in \mathbb{R}_+^{m\times n}$. If $Y = [y_1, y_2, \ldots, y_n] \in \mathbb{R}_+^{m\times n}$, then*

$$\left(\sum_{i=1}^{n} \|y_i\|\right)^{2k} \langle D(D^tD)^k, Y\rangle_f \;\geq\; (\langle D, Y\rangle_f)^{2k+1}. \tag{24}$$

*with equality if and only if $\langle D(D^tD)^k, Y\rangle_f$ is 0, or there exists $\xi > 0$, and $n \times n$ diagonal matrix $\Lambda$ with positive diagonal such that $D^tD\mathcal{P} = \xi^2\mathcal{P}$, and $Y = D\mathcal{P}\Lambda$, where $\mathcal{P}$ is the (0,1)-diagonal matrix such that $\mathcal{P}_{ii} = 1$ if and only if $y_i \neq 0$.*

**Proof.** That the inequality (24) holds follows immediately from Theorem 4. It is the case for equality that requires some explanation. According to Theorem 4 with $p = n$ and $X = \mathrm{I}_n$, equality in (24) occurs if and only if $\langle D(D^tD)^k, Y\rangle_f = 0$, or

$$D\mathcal{P} = \xi Y\mathcal{P}\Lambda, \qquad \text{and} \qquad D^t Y\mathcal{P} = \xi \mathcal{P}\Lambda^{-1}, \tag{25}$$

where $\xi$, $\Lambda$, and $\mathcal{P}$ is as described in the theorem. Let $\hat{\Lambda} = \xi^{-1}\Lambda^{-1}$, and note that $Y\mathcal{P} = Y$. But (25) implies that $D^tD\mathcal{P} = \xi(D^tY\mathcal{P})\Lambda = \xi^2\mathcal{P}(\Lambda^{-1}\Lambda) = \xi^2\mathcal{P}$, and $D\mathcal{P}\hat{\Lambda} = D\mathcal{P}(\xi^{-1}\Lambda^{-1}) = Y\mathcal{P} = Y$. Since these calculations are reversible, the conditions (25) are equivalent to the conditions stated in the theorem with $\Lambda$ replaced by $\hat{\Lambda}$. □

The condition $D^tD = \xi^2\mathrm{I_n}$ is very restrictive, for it asserts that the non-negative matrix $D$ has orthogonal columns. This would mean that no row of $D$ could have more than 1 non-zero entry. If $Y$ has no zero columns, then $\mathcal{P} = \mathrm{I}_n$, and the condition $Y = D\mathcal{P}\Lambda$ reduces to $Y = D\Lambda$, where $\Lambda$ is diagonal with positive diagonal; thus, if $Y$ has no zero columns, then $D$ has no zero columns, and we must have $n \leq m$. If $m = n$, and $Y$ has no columns of zeroes, then the condition for equality in (24), if $\langle D(D^tD)^k, Y\rangle_f \neq 0$, is that $D = \xi Q$, and $Y = Q\hat{\Lambda}$, where $\xi > 0$, $Q$ is a permutation matrix, and $\hat{\Lambda}$ is an $n \times n$ diagonal matrix with positive diagonal.

The appearance of the factor $2k + 1$ in Theorems 2 and 3 suggest that these theorems are simply the odd versions of some more general result. This may be true, however, it is not easy to guess what this more general result is, for the obvious conjecture, namely that there exists an inequality of the form

$$H(X, Y)\,\langle (D^tD)^k X, Y\rangle_f \;\geq\; (\langle DX, Y\rangle_f)^{2k}, \tag{26}$$

where $H$ is some function that is non-negative, and positive when neither $X$ nor $Y$ has a column of zeroes, is simply false. The problem is that it is possible to choose $X$ and $Y$ such that $\langle (D^tD)^k X, Y\rangle_f = 0$, but $\langle DX, Y\rangle_f \neq 0$. To see this consider the special case $m = n = 2$. Let

$$D = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \qquad X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \text{and} \qquad Y = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{27}$$

Then, simple calculations reveal that

$$(D^tD)^k = \begin{bmatrix} 0 & 0 \\ 0 & 2^k \end{bmatrix}, \tag{28}$$

so $\langle (D^tD)^k X, Y\rangle = \langle (D^tD)^k, Y\rangle = 0$. But, $\langle D, Y\rangle = 1$. There is, therefore, no inequality like (26).

Theorem 4 has many corollaries some of which are obtainable without reference to Theorem 4. For example, if we invoke Theorem 4 when $m = n = p$, and $X = Y = \mathrm{I}_n$, we obtain

**Corollary 1.** *If n and k are positive integers, and $D \in \mathbb{R}_+^{n\times n}$, then*

$$n^{2k}\operatorname{trace}\big(D(D^tD)^k\big) \geq \big(\operatorname{trace}(D)\big)^{2k+1}, \tag{29}$$

*with equality if and only if $D = \xi \mathrm{I}_n$ for some $\xi \geq 0$.*

For example, if $D$ were the diagonal matrix $\operatorname{diag}(d_1, d_2, \ldots, d_n)$, where $d_i \geq 0$ for each $i$, then (29) would say that

$$n^{2k}\sum_{i=1}^{n} d_i^{2k+1} \geq \left(\sum_{i=1}^{n} d_i\right)^{2k+1}, \tag{30}$$

but this inequality follows immediately from Hölder's inequality, or from convexity considerations.

If we set $Y = D$ in Theorem 5 we obtain

**Corollary 2.** *If $m$, $n$, and $k$ are positive integers, and $D = [d_1, d_2, \ldots, d_n] \in \mathbb{R}_+^{m\times n}$, then*

$$\left(\sum_{i=1}^{n} \|d_i\|\right)^{2k} \operatorname{trace}\big((D^tD)^{k+1}\big) \geq \big(\operatorname{trace}(D^tD)\big)^{2k+1}, \tag{31}$$

*with equality if and only if $D^tD = \xi^2 \mathrm{I}_n$ for some $\xi \geq 0$.*

Inequality (31) is mysterious in that I do not see how to obtain it independently of Theorem 5. However, if $D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$, where each $d_i \geq 0$, then inequality (31) becomes

$$\left(\sum_{i=1}^{n} d_i\right)^{2k} \left(\sum_{i=1}^{n} d_i^{2k+2}\right) \geq \left(\sum_{i=1}^{n} d_i^2\right)^{2k+1}, \tag{32}$$

which is not difficult to demonstrate using Hölder's inequality.

## References

[1] G.R. Blakley, Prabir Roy, A Hölder type inequality for symmetric matrices with nonnegative entries, Proc. Amer. Math. Soc. 16 (6) (1965) 1244–1245.
[2] Roger A. Horn, Charles R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[3] G. Hardy, J.E. Littlewood, G. Pólya, Inequalities, Cambridge University Press, Cambridge, 1999.
[4] T.H. Pate, Extending the Hölder type inequality of Blakley and Roy to non-symmetric non-square matrices, Tran. Amer. Math. Soc., in press.